UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/896,199	06/29/2001	Eric Cohen-solal	US010324	7568

24737   7590   06/30/2009
PHILIPS INTELLECTUAL PROPERTY & STANDARDS
P.O. BOX 3001
BRIARCLIFF MANOR, NY 10510

EXAMINER
CHANG, KENT WU

ART UNIT	PAPER NUMBER
4145	

MAIL DATE	DELIVERY MODE
06/30/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

UNITED STATES PATENT AND TRADEMARK OFFICE

————————

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

————————

Ex parte ERIC COHEN-SOLAL

————————

Appeal 2008-005770
Application 09/896,199
Technology Center 2600

————————

Decided:[1] June 30, 2009

————————

Before KENNETH W. HAIRSTON, ROBERT E. NAPPI, and BRADLEY
W. BAUMEISTER *Administrative Patent Judges.*

NAPPI, *Administrative Patent Judge.*

DECISION ON APPEAL

———————————

[1] The two-month time period for filing an appeal or commencing a civil
action, as recited in 37 C.F.R. § 1.304, begins to run from the decided date
shown on this page of the decision. The time period does not run from the
Mail Date (paper delivery) or Notification Date (electronic delivery).

This is a decision on appeal under 35 U.S.C. § 134(a) of the rejection of claims 1 through 18 and 20 through 24.

We reverse.

INVENTION

The invention is directed towards a device that allows the picture in picture on a television display to be repositioned or resized. The invention allows a user to use voice commands and gestures to control the picture in picture. See pages 1 and 3 of Appellant's Specification. Claim 1 is reproduced below:

> 1. A video display device comprising:
> a display configured to display a primary image and a picture-in-picture image (PIP) overlaying the primary image; and
> a processor operatively coupled to the display and configured to receive a first video data stream for the primary image, to receive a second video data stream for the PIP, to recognize an audio command related to a PIP display characteristic, the processor, upon recognizing the audio command, activates an image acquisition component that is configured to recognize a user hand gesture related to manipulating the PIP display characteristic, the processor manipulates the PIP display characteristic according to the audio command and the hand gesture.

REFERENCES

Inagaki	US 5,999,214	Dec. 7, 1999
Cox	US 6,154,724	Nov. 28, 2000

VLADIMIR I. PAVLOVIC ET AL., "INTEGRATION OF AUDIO/VISUAL INFORMATION FOR USE IN HUMAN-COMPUTER INTELLIGENT INTERACTION," IMAGE PROCESSING, 1997 PROCEEDINGS IEEE, 121-124.

REJECTIONS AT ISSUE

The Examiner has rejected claims 1 through 18, and 20 through 24[2] under 35 U.S.C. § 103(a) as being unpatentable over Inagaki in view of Pavlovic and Cox. The Examiner's rejection is on page 3 through 10 of the Answer.[3]

ISSUES

Appellant argues on pages 6 through 12 of the Brief[4] that the Examiner's rejection of claims 1 through 18, and 20 through 24 under 35 U.S.C. § 103(a) is in error. Appellant argues that the references do not suggest activating an image recognition component after recognizing an audio command as claimed. Brief 8. Appellant argues that Cox, which the Examiner relies upon for the teaching of voice activation of gesture commands, teaches in response to a voice command, activating a wand which senses gestures and not an image acquisition component as claimed. Brief 9.[5]

Thus, Appellant's contentions present us with the following issue: has Appellant shown that the Examiner erred in finding that the combination of the references teaches a video display device in which an image acquisition

---

component which recognizes hand gestures, is activated by an audio
command as claimed?

PRINCIPLES OF LAW

A rejection based on § 103 must rest upon a factual basis rather than
conjecture or speculation. "Where the legal conclusion [of obviousness] is
not supported by the facts it cannot stand." *In re Warner*, 379 F.2d 1011,
1017 (CCPA 1967). *See also In re Kahn*, 441 F.3d 977, 988.

FINDINGS OF FACT

1. Inagaki teaches a video conference system. Abstract.
2. Inagaki's system includes a movable camera, which has a plurality of
   preset panning positions, at which the camera captures an image of a
   participant. Col. 5, ll. 15-20.
3. Inagaki's system captures a still image of each participant and stores it
   in memory. These still images are displayed in a picture in picture
   display along with one video image of a selected participant. Col. 7,
   ll. 42-61, Figs 8A, 8B, 9A, and 9B.
4. In one embodiment, the participant to be displayed in video is
   determined based upon a voice direction detection unit which
   determines which of the participants in the conference is speaking.
   Based upon that determination the camera pans to the selected
   participant and video of that participant is displayed. Col. 12, ll. 1-25.
5. Pavlovic teaches integration of audio and video in human to computer
   interaction. Pavlovic teaches that individuals prefer to use hand
   gestures in combination with speech when working in a virtual
   environment. Pavlovic, section 1, Introduction, page 121.

6. Pavlovic teaches that the human operator's gestures are captured via a
   video image from a camera that is processed to determine the
   gestures. Pavlovic, section 2.1, Visual Module, Figure 2, page 122.

7. Cox teaches a three dimensional virtual reality system with a gesture
   input interface. Abstract.

8. Cox's system allows the user to utilize voice commands and three
   dimensional spatial tracked gesture inputs. Col. 2, ll. 51-56.

9. Cox teaches that the gestures of the user are observed by using
   magnetically tracked gloves or a hand held tracking device (a wand)
   which is held by the user and has several buttons. Col. 3, l. 9, col. 4,
   ll. 4-7.

10. In one embodiment of Cox's system, menus may be activated by
    voice commands to enable additional functions of the gesture input
    device. Col. 3, ll. 6-10.

ANALYSIS

Appellant's arguments have persuaded us that the Examiner erred in
rejecting claims 1 through 18 and 20 through 24 under 35 U.S.C. § 103(a).
Independent claim 1 recites a processor which receives audio signals and
"upon recognizing the audio command, activates an image acquisition
component that is configured to recognize a user hand gesture related to
manipulating the PIP display characteristic." Independent claims 11, 15, 20,
and 21 similarly recite that there is an activation of an image acquisition
component that determines a gesture of a user upon determining that the
received audio command is recognized. Thus, the scope of the independent
claims includes that the process, whereby an image is acquired to recognize
gestures, is activated by receipt of an audio command.

In rejecting claims 1, 11, 15, 20, and 21, the Examiner finds that Inagaki teaches a video display device with a picture in picture image overlaying a primary image and that Inagaki teaches that the image is controlled from voice indication. Answer 3 and 4. Further, the Examiner finds that Inagaki does not teach upon recognizing the audio command activating an image acquisition component. Answer 3 and 4. We find ample evidence to support these findings. Facts 1-4. The Examiner finds that Pavlovic demonstrates a system that uses audio commands and related gestures to control a graphical object. Answer 4. Similarly, we find ample facts to support these findings. Fact 5. The Examiner finds that it is an obvious design choice to choose whether to enter a voice command first and then a gesture command. The Examiner cites Cox as evidence of the design choice. Answer 5. We disagree with the Examiner's reasoning.

Initially, we note that the claims are narrower than the voice command having to be received before the gesture as stated by the Examiner. Rather, as discussed above the claim recites that the image acquisition which interprets the video is actuated by the voice command. The teachings of Cox are directed to a system which allows a user to use voice commands and gestures in a three dimensional data structure. Facts 7 and 8. The user's gestures are captured by a magnetically tracked glove or a hand held tracking device (referred to as a "wand"). Fact 9. Cox does not teach that the gesture input device is achieved by image acquisition (although Pavlovic does, Fact 6). Further, Cox does not teach that the acquisition of the gesture from the gesture input device is activated by a voice command. Rather, Cox teaches that menus in the image system are voice activated which enables additional functions to be input by gestures. Fact 10. The Examiner's rationale on page 5 of the Answer directed to a gesture control of menus is

similarly not persuasive as it discusses the actuation of a command that
requires cursor control and not actuation of an image acquisition component
as claimed. Thus, we do not find that Cox provides evidence or suggests
that the teachings of Inagaki and Pavlovic should be modified such that the
input device which recognizes gestures is activated by receipt of an audio
command. Accordingly, we will not sustain the Examiner's rejection of
claims 1 through 18, and 20 through 24 under 35 U.S.C. § 103(a).

ORDER

The decision of the Examiner to reject claims 1 through 18 and 20
through 24 is reversed.

<div align="center"><u>REVERSED</u></div>

ELD

PHILIPS INTELLECTUAL PROPERTY & STANDARDS
P.O. BOX 3001
BRIARCLIFF MANOR, NY 10510